Nucl Eng Technol 47 (2015) 176–186

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.journals.elsevier.com/nuclear-engineering-and-technology/

Original Article

SEMISUPERVISED CLASSIFICATION FOR FAULT DIAGNOSIS IN NUCLEAR POWER PLANTS

*JIANPING MA and JIN JIANG**

Department of Electrical and Computer Engineering, University of Western Ontario, 1151 Richmond Street North, London, Ontario N6A5B9, Canada

ARTICLE INFO

Article history:
Received 10 October 2014
Received in revised form
20 December 2014
Accepted 22 December 2014
Available online 21 January 2015

Keywords:
Fault diagnosis
Nuclear power plant
Semisupervised classification

ABSTRACT

Pattern classifications have become important tools for fault diagnosis in nuclear power plants (NPP). However, it is often difficult to obtain training data under fault conditions to train a supervised classification model. By contrast, normal plant operating data can be easily made available through increased deployment of supervisory, control, and data acquisition systems. Such data can also be used to train classification models to improve the performance of fault diagnosis scheme.

In this paper, a fault diagnosis scheme based on semisupervised classification (SSC) scheme is developed. In this scheme, new measurements collected from the plant are integrated with data observed under fault conditions to train the SSC models. The trained models are subsequently applied to new measurements for fault diagnosis. In comparison with supervised classifiers, the proposed scheme requires significantly fewer data collected under fault conditions to train the classifier.

The developed scheme has been validated using different fault scenarios on a desktop NPP simulator as well as on a physical NPP simulator using a graph-based SSC algorithm. All the considered faults have been successfully diagnosed.

The results have demonstrated that SSC is a promising tool for fault diagnosis in NPPs.

1. Introduction

Safety and availability of a nuclear power plant (NPP) can be adversely affected by various component faults. It is important to diagnose potential faults early enough so that a minor fault may not develop into potentially disastrous consequences. Pattern classification has become an important tool for early fault diagnosis in NPPs and other industries [1–11]. The processes in a NPP are governed by physical laws such as balances of mass, energy, and momentum. When process

* *Corresponding author.*
E-mail address: jjiang@eng.uwo.ca (J. Jiang).

http://dx.doi.org/10.1016/j.net.2014.12.005
1738-5733/Copyright © 2015, Published by Elsevier Korea LLC on behalf of Korean Nuclear Society.
Special Issue on ISOFIC/ISSNP2014.

faults have happened, these balances are altered. For different types of faults, the system reacts differently. As a result, different patterns of system behaviors can be observed. Therefore, with *a priori* knowledge of the system behaviors under specific fault, and with a proper selection of a set of measurements, the faults could be uniquely identified through pattern classification means.

In case of NPP applications, one of the problems is the lack of adequate training data to match faults and their corresponding symptoms. This makes pattern based fault classification particularly difficult in NPPs [12,13]. The majority of pattern classification based fault diagnosis applications use so-called supervised pattern classification models, where training data with known fault classes are used to train a classifier first. The classifier subsequently processes new measurements to diagnose potential faults by matching the patterns against the measurement data. However, reliable data under a specific fault condition are rare in NPP to train a classifier. One of the reasons is that one cannot simply inject real faults into an NPP system for the purpose of collecting training data. Another reason is that most of the critical system components in a NPP have gone through rigorous qualification process, inspection, and they are often of high quality and extremely reliable. Although there are some databases to describe fault conditions of some components, they may not coincide exactly with the fault classes considered by the diagnostic classifier. Use of an NPP simulator can be another way to generate training data, but there are inevitably differences between the simulator responses and the real plant responses due to modeling errors. Scarcity of reliable data under fault conditions can skew boundaries of a classifier, and subsequently lead to false classification. Therefore, there is a need to develop a fault diagnosis scheme, which relies less on the known fault patterns. A fault diagnosis scheme based on semisupervised classification (SSC) is developed in this paper to address the above issue.

In the terminology of classification, the data set under specific conditions are known as labeled data, whereas the data set with unknown conditions are referred to as unlabeled data. The objective of a classifier is to uncover specific conditions (i.e., faults) based on these unlabeled data (i.e., new measurements), and assign appropriate labels (i.e., fault types) to them accordingly. In the proposed scheme, both labeled and unlabeled data are used to train the classifier. Hence, it is called SSC. Once trained, the classifier can then be used to process the new measurements and perform fault diagnosis. There are two principles that most SSCs are based on: (1) *clustering assumption*, which states that nearby data points probably belong to the same class; and (2) *manifold assumption*, which says that data points on the same manifold structure are likely to be in the same class [14,15]. An SSC model can achieve superior performance because the classifier can be designed to avoid cutting through the high-density regions or the manifolds when processing the unlabeled data. In addition, a higher degree of uncertainties in the labeled data can be tolerated. In the case of physical systems in an NPP, correlations often exist among different measurements due to their physical and functional couplings. Therefore, data collected under the same fault condition tend to fall in the same high-density region or on the same manifold structure. For these reasons, SSC has become a promising fault diagnosis tool for applications in NPPs.

In the proposed fault diagnosis scheme, type of faults considered by the diagnostic system has to be defined first. Sensors that can be used to collect data for diagnostic purpose are also selected. Labeled data under these fault conditions are then generated through various means, such as by simulation, experiments on scaled physical mock-ups, or experience on the cause–effects associated with system components. If a fault is detected, measurements are collected and treated as unlabeled data. These data are then integrated with the available labeled data samples to train an SSC model. Once the model is trained, it will subsequently be used to classify/assign the most appropriate labels for the unlabeled inputs. In other words, they diagnose the underlying fault conditions embedded in the collected data; hence, the fault diagnosis task can be accomplished. The scheme has been validated using a desktop simulator of a Canada deuterium uranium (CANDU) NPP and a physical simulator of an NPP known as a NPP control test facility (NPCTF) that is essentially a simplified physical component based NPP simulator for instrumentation and control purposes. Three types of faults are considered in the CANDU simulator and six types of faults are on the NPCTF. Classification results have shown that all faults can be successfully diagnosed, even though the labeled data used contain a considerable amount of uncertainties, and the size of the labeled data is significantly smaller than typically required in a supervised classification scheme.

2. Materials and methods

2.1. *Pattern classification for fault diagnosis*

Suppose k faults can occur in a system and m sensors are used for data collection to diagnose them. The sensor outputs sampled at a time point t are organized into a vector $x_t \in R^{1\times m}$. Denote the hypothesis for the i-th fault as Hi, and hence, the fault is represented by a class label $y = i$. Normal operation is denoted as $H0$ and the class label is assigned as $y = 0$. In addition, it is assumed that a set of training data are available and denoted as $D^l = (X^l, Y^l) = [(x_i^l, y_i)]_{i=1,\ldots,nl}$, where nl is the total number of labeled data, X^l contains all the inputs and Y^l contains the class labels associated with X^l. The training data are selected to capture the uniqueness in sensor responses under different faults. When a fault is detected, a set of measurements $X^u = (x_i^u)_{i=1,\ldots,nu}$ is acquired, where nu is the number of unlabeled new measurements. The objective of a fault diagnosis system is to determine which fault has occurred by identifying the closest match of the underlying patterns between the new measurements and the training data model. This is essentially a pattern classification problem of assigning discrete labels to prove or disapprove the fault hypotheses, i.e., $y \in (0, 1, 2, \ldots, k)$ based on a set of new data X^u, given a set of training data D^l. In this paper, X^l and X^u are known as the labeled and the unlabeled data sets, respectively.

To implement a pattern classification system, a classification function $g_i(y = i|x)$ has to be defined for each class, which characterizes the match of the inputs x to the class $y = i$. The classification process can be expressed as:

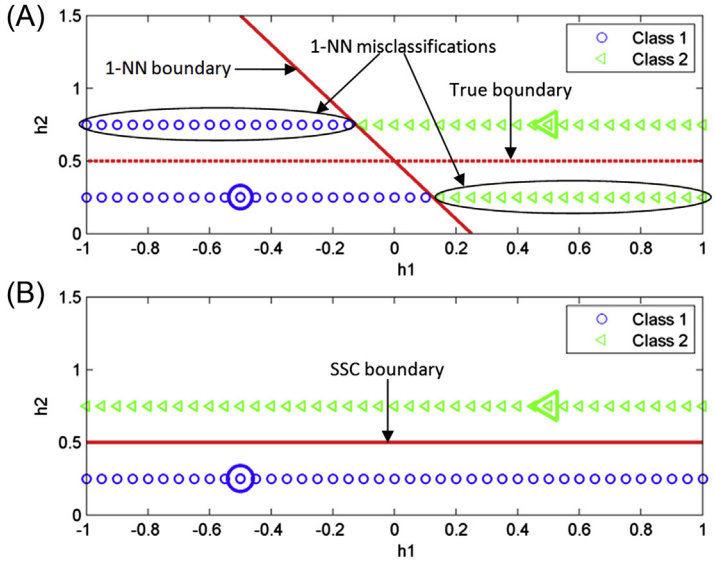

Fig. 1 – Supervised classification. Classification is based on (A) 1 – nearest neighbor, and (B) semisupervised.

$$y = \underset{i,i=1,\ldots,k}{\operatorname{argmax}}\, g_i(y = i|x) \tag{1}$$

From a probabilistic point of view, a classification function can be modeled from a generative approach to learn the joint probability distribution of the inputs x and the class label y. The classification function can also be modeled from a discriminative approach to find a direct map from the inputs to the class label [16–18]. Taking a Bayes classifier as an example, the posterior probability is used as the classification function as:

$$g_i(y = i|x) = p(y = i|x) \tag{2}$$

Through the Bayes rule, Equation 2 is equivalent to:

$$g_i(y = i|x) \propto p(x|y = i)p(y = i) \tag{3}$$

where $p(x|y = i)$ is the *a priori* distribution of the measurements for class $y = i$ and $p(y = i)$ is the probability distribution of class $y = i$.

The *a priori* distribution $p(x|y = i)$ must be estimated from the training data, which can be obtained using various techniques, e.g., maximum likelihood estimation (MLE). In MLE, it is assumed that $p(x|y = i, \theta)$ belongs to certain distribution and θ represents a set of parameters in the selected distribution. With a supervised pattern classification model, MLE estimates the value of θ by maximizing the log likelihood of $p(D^l|\theta)$[16–18] as:

$$\log p(D^l|\theta) = \log \prod_{j=1}^{nl} \left(p\left(x_j, y_j|\theta\right)\right) = \sum_{j=1}^{nl} \log p\left(y_j|\theta\right) p\left(x_j|y_j, \theta\right) \tag{4}$$

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmax}} \left(\log p\left(D^l|\theta\right)\right) \tag{5}$$

If size of the training data set D^l is small, the training data may not cover all the scenarios considered for the classification tasks. As a result, errors may occur in the estimated parameters $\widehat{\theta}$. Consequently, performance of the classifier can be affected. In the case when the training data set is small, but unlabeled data are abundant, i.e., $nu >> nl$, SSC models can enhance the performance of the classifier. In SSC, both labeled and unlabeled data are used to train the classifier. Taking the Bayes model for example again, the model parameters θ are obtained by maximizing the joint log likelihood defined on both the labeled training data D^l as well as the unlabeled instances X^u [17], i.e.,

$$\begin{aligned} \log p\left[\left(D^l, X^u\right)|\theta\right] &= \log \prod_{j=1}^{nl} p\left[\left(x_j^l, y_j\right)|\theta\right] \prod_{i=1}^{nu} p\left(x_i^u|\theta\right) \\ &= \sum_{j=1}^{nl} \log p\left(y_j|\theta\right) p\left(x_j|y_j, \theta\right) + \sum_{i=1}^{nu} \log p\left(x_i|\theta\right) \end{aligned} \tag{6}$$

When Equation 6 is compared to Equation 4, it can be seen that information contained in the unlabeled data also contributes to the MLE estimation of the distribution parameters. The $\widehat{\theta}$ obtained from Equation 6 will be different from those obtained based solely on the labeled data as in Equation 4. Because the unlabeled data may cover additional regions that may not be covered by the limited amount of training data, it can potentially lead to enhanced performance. Note that Equations 4 and 6 are used to illustrate both the connections and the differences between a supervised and a semisupervised pattern classification. In fact, the unlabeled data can be integrated into a pattern classification model through various means as summarized in review papers and books [14,19–21].

Supervised and semisupervised classifications are compared in Fig. 1 using a simple two-class example, where the input is a vector of two variables $x = (h_1, h_2)$ and the classes are designed as $y = 1|h_2 = 0.25$ and $y = 2|h_2 = 0.75$. The boundary of the desirable classifier is at $h_2 = 0.5$, as shown in Fig. 1A using the red broken line. Suppose two training data samples are available as $D^l = \{d_1, d_2\} = \{[(-0.5, 0.25), 1], [(0.5, 0.75), 2]\}$, which are shown in Fig. 1 as the large size blue circle and green triangle, respectively. If a supervised 1 – nearest neighbor (1–NN) classifier [22] is trained based on d_1 and d_2, the classification boundary is essentially a line vertical to the

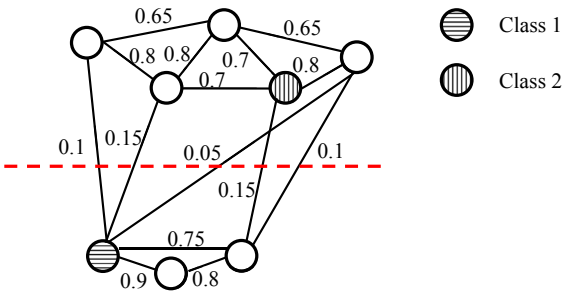

Fig. 2 – A data graph example.

line connecting d_1 and d_2. This boundary is shown as the red unbroken line in Fig. 1A, which is considerably different from the truth. Some new unlabeled data in the range of $h_1 \in [-1, 1]$ are generated and classified by the 1–NN classifier. The results are shown in Fig. 1A, where data classified to the two classes are represented using the smaller size circles and triangles, respectively. It is observed that a large portion of the unlabeled data has been misclassified by the 1–NN classifier. The misclassified data are pointed out using the two ovals. The same unlabeled data and training data are then processed by a SSC model and the results are presented in Fig. 1B. It can be seen that all unlabeled data has been correctly classified. It represents a substantial improvement as compared to the results in Fig. 1A.

It should be noted that, even though it has been demonstrated that SSC can produce superior performance in numerous studies, it should not be taken for granted that indiscriminate inclusion of unlabeled data in a pattern classification model will lead to improved performance [23,24]. For instance, it has been shown [24], that it is important to ensure that there exists a truly nontrivial relationship between distribution of the unlabeled data and the class labels.

2.2. A graph-based SSC model

Several SSC algorithms have been developed recently [14,19–21]. A graph-based SSC algorithm [25,26] was chosen for the current study due to its proven classification performance and the ability to perform multiclass classification by a single classifier. This algorithm can be considered as a combination of spectral clustering [27–29] and label propagation [30] on a data graph. First, a graph $G(V, E)$ is constructed for a data matrix $X = \{x_i\}_{i=1,\ldots,n}$, where the vertices V are the data points in X and E are edges connecting the vertices. A weight w_{ij} is assigned to the edge connecting vertices i and j, which can be computed as:

$$w_{ij} = \exp\left[-\frac{(x_i - x_j)^2}{2\sigma^2}\right], i \neq j \tag{7}$$

where σ is a constant. $w_{ij} = 0$ if $i = j$. An affinity matrix W is formed with weights for all the data points.

A diagonal degree matrix D can then be defined as:

$$D = diag(D_{ii}) = diag\left(\sum_j w_{ij}\right) \tag{8}$$

Elements in W quantify similarities between two data samples and elements in D quantify the total weights for every data sample.

An example data graph is shown in Fig. 2 to illustrate the concept. The circles are the vertices, representing the input data points. The two filled circles represent labeled data for two classes as shown by the legend. The lines connecting the vertices are the edges. The numbers beside the edges are the weights. In a completed data graph, all pairs of data points are supposed to be connected. However, only some of the connections are shown in Fig. 2, so that the graph does not look overly crowded. A larger weight means that the two data points are more similar or closer to each other. When a label propagation process is performed on the graph, a larger weight means that the likelihood of receiving label information from the other data point is higher. Intuitively, it appears to be reasonable to partition the data points in Fig. 2 in such a way that data below the broken line belong to Class 1 and data above the broken line belong to Class 2.

Spectral clustering is formulated as a graph cut problem so that the edges between different clusters have smaller weights and edges within the same cluster have larger weights. Minimizing the graph cut functions becomes an eigenvalue problem [29] as:

$$Lu = \lambda Du \tag{9}$$

where u is an eigenvector, λ is an eigenvalue and L takes the form:

$$L = D^{-1/2} W D^{-1/2} \tag{10}$$

Classification by SSC is achieved by integrating spectral clustering with some labeled data. Create the input data matrix $X = \begin{bmatrix} X^l \\ X^u \end{bmatrix} \in R^{n \times m}$, where $n = nl + nu$ is the total number of labeled and unlabeled data. Define a classification matrix $F \in R^{n \times (k+1)}$. Each row in F represents the similarities of one data sample to all the classes to be assigned. The label of a data point x_i can then be obtained by $y_i = \underset{j,j=1,\ldots,k+1}{\operatorname{argmax}} F_{ij}$.

In the algorithm in [25,26], every data point in X spreads its label information to the neighbor points with the matrix L specifying the likelihoods. The label spread process iterates until this process is stable. It is more likely for a labeled datum to propagate its label information to data in the same high-density region or on the same manifold structure. Referring to the example in Fig. 2, the three data points below the broken line are more likely to propagate their label information to each other than to the rest of the graph. When a convergent stable state is achieved, the unlabeled data points are more likely to possess label information from the labeled node for Class 1. It is the same situation for the data points above the broken line.

Mathematically, the classification matrix F is obtained iteratively as [25]:

$$F(t+1) = \alpha L F(t) + (1 - \alpha) Y \tag{11}$$

where α is a parameter between 0 and 1 that controls the relative importance of the two terms on the right side. It can be seen from Equation 11 that the classification is based on information received from its neighbors (the 1st term) and

also the initial information (the 2nd term). The convergent value of F can be obtained as [25,26]:

$$F^* = (1-\alpha)(I-\alpha L)^{-1}Y \tag{12}$$

and the class label of an unlabeled data can be obtained as:

$$y_i = \underset{j,j=1,\ldots,k+1}{\operatorname{argmax}} F^*_{ij} \tag{13}$$

A complete SSC algorithm can be summarized as follows: (1) collect the labeled data and the unlabeled data into a data matrix X and initiate the label matrix Y; (2) form the weight matrix W as $w_{ij} = \exp[-(x_i - x_j)^2/2\sigma^2], i \neq j$ and $w_{ii} = 0$; (3) construct the diagonal degree matrix D as $D = diag(D_{ii}) = diag(\sum_j w_{ij})$; (4) compute $L = D^{-1/2}WD^{-1/2}$; (5) obtain the classification matrix F as $F = (1-\alpha)(I-\alpha L)^{-1}Y$; and (6) label a data point as $y_i = \underset{j,j=1,\ldots,k+1}{\operatorname{argmax}} F^*_{ij}$.

2.3. A fault diagnosis scheme based on SSC

In this section, a fault diagnosis scheme based on SSC is developed. The logical flow is summarized in Fig. 3. The entire process can be divided into two parts: the first part (Steps 1 and 2) is for off-line training and data preparation; and the second part (Steps 3–7) is for on-line fault diagnosis. At Step 1, the fault types considered are determined first. One can use techniques such as failure mode and effect analysis for this purpose. Class labels are assigned to the appropriate fault classes. Inputs to the SSC model are also selected so that all the faults can be distinguished. Desirable sensor placement methods can be used [31,32] to select the most suitable sensors for this purpose. Labeled training data are then generated in Step 2. For this purpose, measured data with real faults, historical operation data, operating experience, and numerical simulations can all be considered to generate the labeled data. At Step 3, new measurements are acquired during system operation. Step 4 is a process that detects potential faults based on the new measurements. If a fault is detected, the new measurements in sequel will be logged as the unlabeled data, which is known as Step 5. In Step 6, both the labeled and the unlabeled data are processed by a SSC model to determine the most appropriate class labels for the unlabeled new measurements. Based on the fault class designation, at Step 7, the class labels are assigned to identify whether the system is normal or, if not, what type of fault has probably occurred. Steps 3–7 are intended for on-line fault detection and diagnosis.

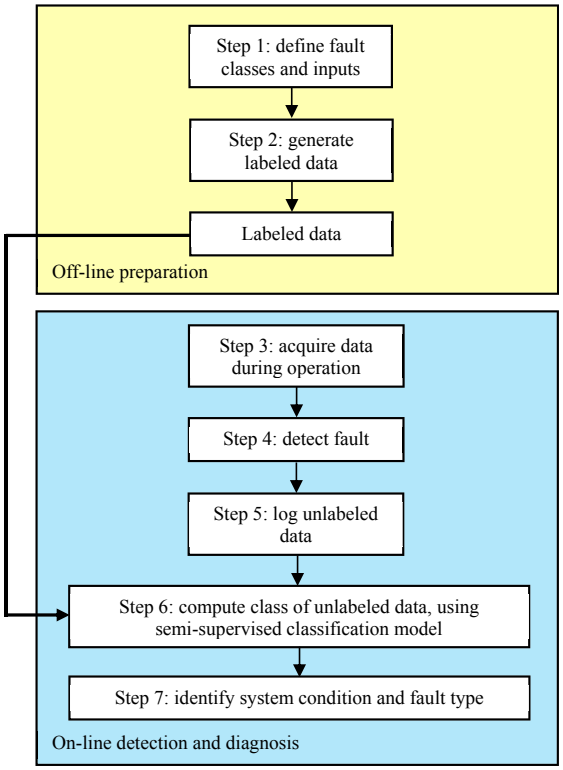

Fig. 3 – Procedures of the proposed fault diagnosis scheme.

3. Results

The developed scheme has been validated by two case studies. One is using a desktop CANDU NPP simulator and the other one is on the NPCTF. In these case studies, a data matrix X for the graph-based SSC algorithm has been constructed as:

$$X = \begin{bmatrix} X_{n0\times m} \\ X_{n1\times m} \\ \ldots \\ X_{nk\times m} \\ X_{nu\times m} \end{bmatrix} \tag{14}$$

where nk represents the size of the labeled data for class $y = k$ and nu is the number of unlabeled data. The classification matrix F is initiated as shown in Equation 15. A zero matrix of the size $nu \times (k+1)$ is assigned for the unlabeled data. For a labeled data point, the i-th entry of the corresponding row vector is set as unity, and the remaining ones are set to zero. For $H0$, the last entry is set to unity:

$$Y = \begin{bmatrix} [0,0,0,\ldots,1]_{n0\times(k+1)} \\ [1,0,0,\ldots,0]_{n1\times(k+1)} \\ [0,1,0,\ldots,0]_{n2\times(k+1)} \\ \ldots \\ [0,0,\ldots,1,0]_{nk\times(k+1)} \\ [0,0,\ldots,0,0]_{nu\times(k+1)} \end{bmatrix} \tag{15}$$

3.1. Case study using a desktop CANDU simulator

To generate data under different fault conditions, a high fidelity desktop NPP simulator of a CANDU plant is used. In the case study, three faults in the pressurizer, feedwater, and main steam systems have been simulated. These faults can affect various subsystems in the plant, which can result with similar symptoms in different system variables. The first simulated fault is a spurious opening of one pressurizer steam bleed valve. The steam bleed valve is used to reduce the pressure in the heat transport system by bleeding the steam out. The second fault is a spurious closure of one boiler level control valve that is used to control the steam generator level by regulating the feedwater flow rate. In addition, a low position fault is simulated to the backup boiler level control valve. The third fault is spurious opening of one main steam safety valve, which is used to protect the steam lines. The following seven process variables are used as the model

Table 1 – Validation data sets using a Canada deuterium uranium nuclear power plant simulator.

Data set	Simulation condition	Fault hypothesis	Class label
1	Normal operation	*H0*	0
2	1 boiler level control valve fail close; the backup boiler level control valve has a 30% position fault	*H1*	1
3	1 steam bleed valve stuck open at 80%	*H2*	2
4	Main steam safety valve fail open	*H3*	3

inputs to diagnose the faults: (1) gross reactor power (%); (2) reactor outlet header pressure (kPa); (3) steam generator (SG) level (m); (4) SG pressure (kPa); (5) feedwater flow (kg/s); (6) steam flow from SG to the balance header (kg/s); and (7) balance header pressure (kPa). As summarized in Table 1, four sets of data are collected from the simulator. The simulations are all initiated with full-power steady state condition. No fault is inserted for the first simulation. Each of the other three simulations has one different fault simulated. To generate labeled training data, one out of every 20 data points in the first data set is used as the labeled training data for *H0*. Four data points are selected from each of the other data sets as the labeled data points for *H1*, *H2*, and *H3*. The rest of the data in all data sets are treated as unlabeled data.

Classification results are illustrated in Fig. 4 using three variables, i.e., SG pressure, SG level, and reactor outlet header pressure. The labeled data is shown in Fig. 4 as the larger size symbols and the unlabeled data is represented using the smaller size symbols. Different shapes and colors are used to distinguish the classes. It can be observed that the unlabeled data are all classified to their respective fault classes based on measurements collected under different fault conditions. The class labels estimated for the four data sets as a function of time are illustrated in Fig. 5. It can be seen in Fig. 5A that all data points with the normal operation are correctly classified as hypothesis *H0*. For the data sets with faults as shown in Fig. 5C and D, the data are initially classified as normal because the data set also contains the measurements under full-power steady state

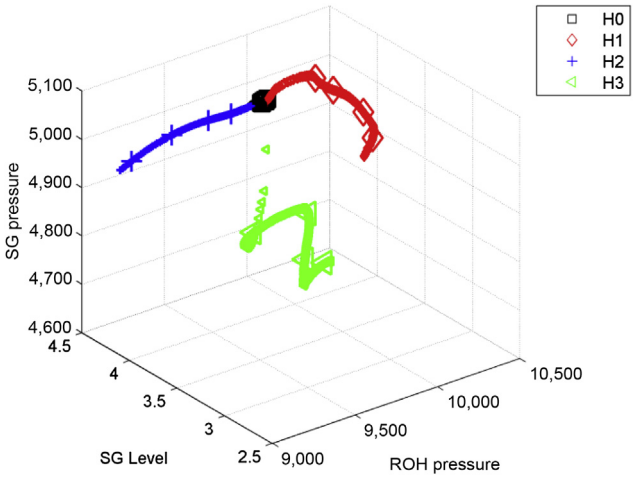

Fig. 4 – Classification results for faults in a Canada deuterium uranium nuclear power plant.

condition at the beginning. The faults are all correctly classified shortly after as more data are received. As demonstrated by the results in Figs. 4 and 5, the SSC model is able to distinguish all three faults with only a small number of labeled data.

3.2. *Case study using a physical simulator for NPP instrumentation and control studies*

The developed fault diagnosis scheme has also been validated using experimental data collected on the NPCTF, which is a physical simulator for NPP instrumentation and control studies. A picture of the system is shown in Fig. 6 and the physical process chosen for this study is also illustrated in Fig. 7. The system mainly consists of a pump, a heater, a heat exchanger, a chiller unit, and a pressurizer tank. The pump drives the primary working fluid (water) flow and CV-1 controls the flow rate. The water is heated up by the heater, which simulates the function of a reactor. The heated water is cooled in the heat exchanger using chilled water from the chiller, simulating the heat sink function of a SG. The pressurizer tank regulates the system pressure by feeding in the compressed air through CV-3 or bleeding the air out through CV-4. Two manual valves, FV-1 and FV-2, are used to simulate pipe breaks or flow blockage. Valve openings can be set manually in manual control mode or regulated by a controller in automatic control mode. As shown in Fig. 7, the process measurements include flow F1, pressure P1 and P2, temperature T1 and T2, and level L1.

As summarized in Table 2, six faults are deliberately created on the NPCTF for this case study. They are: (1) pipe break by opening FV-1; (2) CV-5 spurious open; (3) reactor overpower by excessive heating; (4) CV-2 spurious close; (5) CV-1 spurious close; and (6) flow blockage by closing FV-2. Seven sets of data are collected from the NPCTF with a sampling interval of one second. The first data set is obtained when the system is under normal steady state operation. Each of the other six sets of data is acquired with a different fault. The faults are created independently when the system is under the normal steady state operating condition. Statistics of the process measurements under the normal operation are obtained and presented in Table 3. Based on the data in Table 3 and the NPCTF operating characteristics, a fault is considered as being detected if any sensor output exceeds the thresholds listed in Table 4.

Two different ways are used to generate the labeled data in this case study. The first one is to generate the labeled data based on dynamic relations among different variables in the NPCTF. The responses of the process variables to a fault can be identified from the dynamic relations, thus data collected

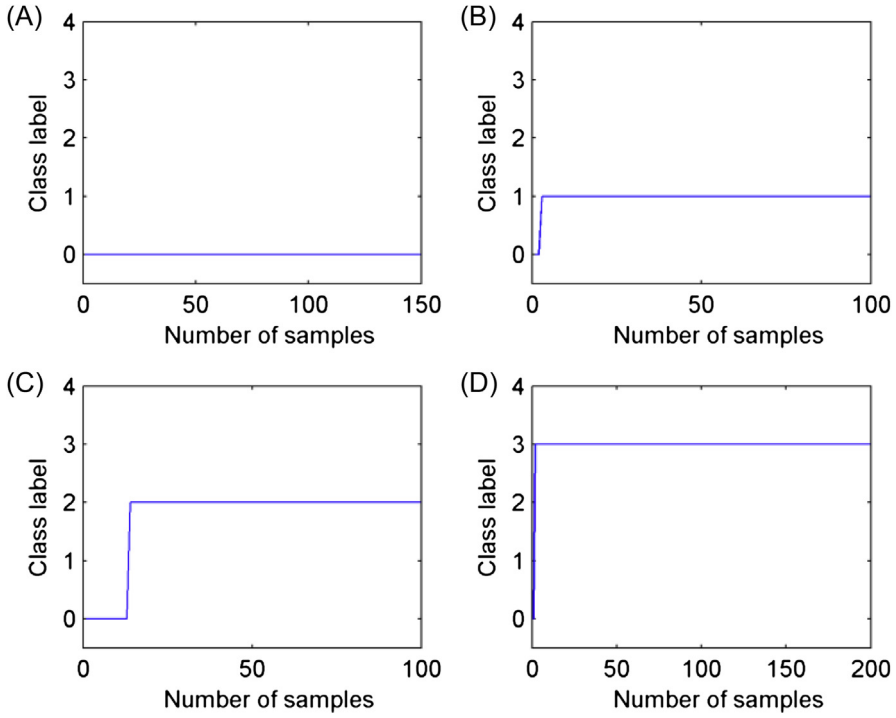

Fig. 5 – Class labels for the Canada deuterium uranium nuclear power plant data sets: (A) normal operation *H0*; (B) boiler level control valve fault *H1*; (C) steam bleed valve fault *H2*; and (D) main steam safety valve fault *H3*.

from the sensors can be approximated. It is a way to incorporate previous experiences and *a priori* knowledge to assist the pattern classification task. This method is used to generate the training data for the faults related to FV-1 (*H1*) and FV-2 (*H6*), as summarized in Table 5. To use the case of FV-2 as an example, when FV-2 is partially closed, the water flow in the primary loop is restrained, thus the flow rate F1 will drop and the pressure P1 will increase. In addition, the heater outlet temperature T2 will increase due to the reduced flow. No substantial changes in other

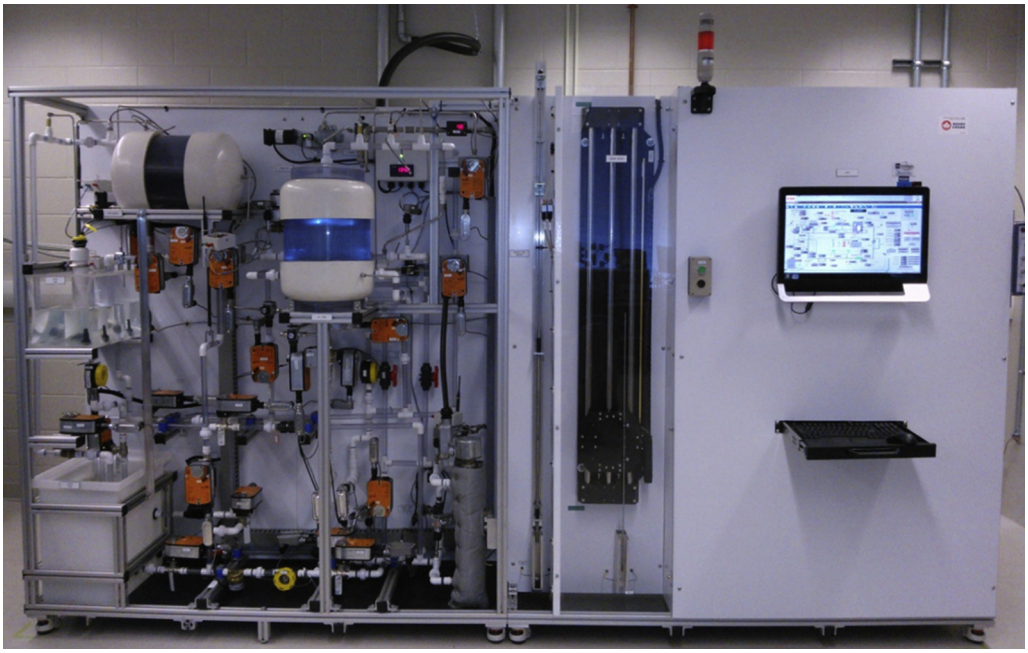

Fig. 6 – Nuclear power plant control test facility.

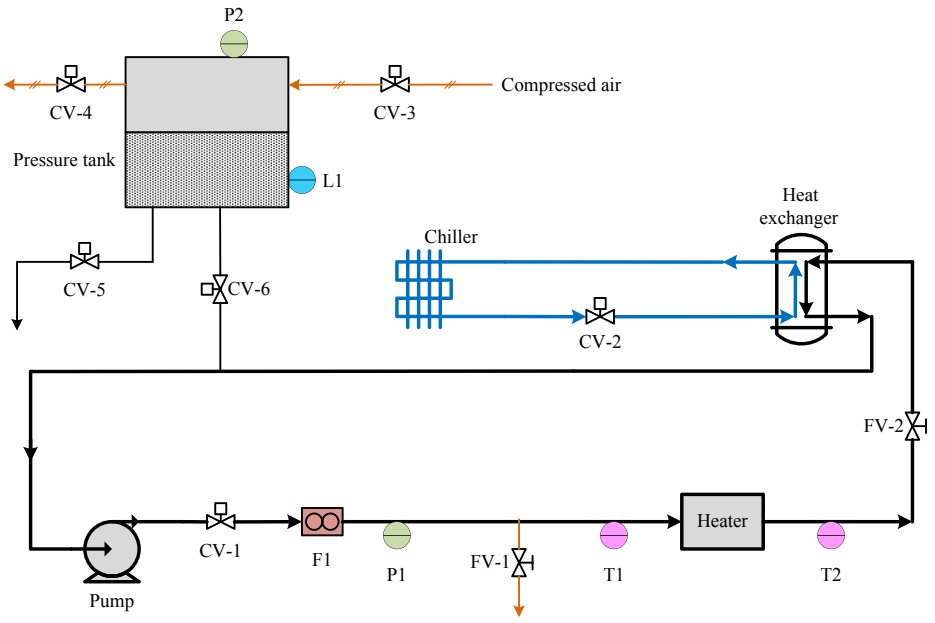

Fig. 7 – Process diagram of the selected nuclear power plant control test facility loops.

Table 2 – Summary of case studies on nuclear power plant control test facility (NPCTF).

Data set	NPCTF operation	Fault hypothesis	Class label
1	Normal	*H0*	0
2	Manually open FV-1 to a high position	*H1*	1
3	Open CV-5 to a high opening in manual mode	*H2*	2
4	Set heater power high in manual mode	*H3*	3
5	Set CV-2 at a low opening in manual mode	*H4*	4
6	Set CV-1 at a low opening in manual mode	*H5*	5
7	Manually close FV-2 to a low opening	*H6*	6

measurements are expected. Based on this knowledge and the normal values presented in Table 3, six labeled data points are generated as follows: F1 has a range of 5.5–3.0; P1 has a range of 9.5–12.0; T2 has a range of 32.5–35.0; and values of other variables are selected arbitrarily within the normal operating ranges. Six labeled data are also generated for *H1* in a similar fashion as summarized in Table 5 (1 psi = 6.89 kPa). As expected, the training data generated this way will probably contain considerable uncertainties. If more accurate models are available to simulate the dynamic relations, they can also be used to form the labeled data.

The second way to obtain the labeled data for the SSC model is to pick some samples from the test data set as the labeled data. For the normal hypothesis *H0*, every 30 data points of the first data set are used as the labeled data. It is a valid method to produce the labeled data because the system can repeat the normal steady state operations to generate data. For the faults related to CV-5, heater, CV-2, and CV-1 (*H2*-*H5*), only two data samples are picked from the corresponding test data sets as the labeled training data. The reason that the labeled data can be produced in such a way is that, for some fault scenarios, the actual effects can be physically tested on NPCTF or emulated. The frequency and the severity of a fault

Table 3 – Statistics of nuclear power plant control test facility normal operation data.

Variable and unit	Minimum	Maximum	Average	Standard deviation
F1 (l/m)	6.26	6.61	6.44	0.09
P1 (psi)	9.01	9.13	9.07	0.02
T1 (°C)	18.29	21.48	19.89	1.04
T2 (°C)	28.15	31.06	29.61	0.97
P2 (psi)	6.04	6.03	6.03	0.005
L1 (%)	39.84	39.90	39.87	0.01

Table 4 – Fault detection thresholds for nuclear power plant control test facility.

Variable and unit	Lower bound	Higher bound
F1 (l/m)	6.1	6.8
P1 (psi)	8.9	9.2
T1 (°C)	17.5	22.0
T2 (°C)	27.5	31.5
P2 (psi)	5.8	6.2
L1 (%)	39.0	41.0

Table 5 – Labeled data for fault *H6* and *H1*.

Labeled data		F1 (l/m)	L1 (%)	P1 (psi)	P2 (psi)	T1 (°C)	T2 (°C)
Fault hypothesis *H6*	1	5.5	40.2	9.5	6.0	19.0	32.5
	2	5.0	40.0	10.0	6.1	18.5	33.0
	3	4.5	39.5	10.5	5.9	20.0	33.5
	4	4.0	40.1	11.0	6.0	19.5	34.0
	5	3.5	40.2	11.5	6.2	20.0	34.5
	6	3.0	40.1	12.0	6.1	18.0	35.0
Fault hypothesis *H1*	1	7.1	38.5	8.5	5.5	21.0	32.5
	2	7.2	37.5	8.0	5.0	20.5	33.5
	3	7.4	36.0	7.5	4.5	21.0	34.5
	4	7.3	34.0	7.0	4.0	20.0	35.5
	5	7.2	33.0	6.5	3.5	19.5	36.5
	6	7.4	32.0	6.0	3.0	21.0	37.5

injection may be very limited in order to control the induced stress within the safety boundaries, but it provides a viable way to understand the underlying dynamic relationships of the real process.

The sensor outputs for the seven sets of experiment data are shown in Fig. 8A and class labels estimated by the SSC model for the seven data sets are shown in Fig. 8B. Different sets of data are separated by the vertical lines in Fig. 8. The horizontal axis of the plots in Fig. 8 is the number of samples which is equivalent to time in seconds. It can be seen (Fig. 8B) that normal operation data are all correctly classified. For the fault of FV-1 open, it is observed (Fig. 8B) that all the data are correctly classified to fault hypothesis *H1* except the first few data points, which are mistakenly classified as normal due to the dominance of the normal operating data at the beginning. Fault detection of FV-1 is triggered by low pressure which responds fast to the opening of FV-1. However, changes in other parameters, such as temperature, could be slower. Therefore, the unique symptoms associated with this fault have not yet fully developed at the beginning, which leads to the misclassifications. The case for the spurious opening of CV-5 is similar to FV-1. The data at the beginning conclude a normal operation, but correct classification is achieved subsequently. The number of misclassified data is larger than the case of FV-1 because the physical movement of CV-5 is slow. Therefore, it takes longer for the fault to propagate. The classification results for the other four faults (heater, CV-2, CV-1, and FV-2) are also shown in Fig. 8B. The faults in the test cases are all classified satisfactorily.

Classification results of the seven data sets are further illustrated in Fig. 9, using three variables F1, L1, and T2 in the first plot on the top and using F1, P1, and T1 in the second plot at the bottom. The labeled data points are shown as the larger size symbols and the unlabeled data are shown as the smaller size symbols. In addition, different classes are represented using different symbol shapes and colors. For the data set with FV-1 fault (*H1*), pink circles are used. When the first plot in Fig. 9 is examined, it can be seen that the labeled data, shown as the larger size circles, has considerable uncertainties as compared to the measured data, shown as the smaller size circles. The reason is that the six labeled data points are generated based on coarse analysis of the relations among the process variables without detailed mathematical models or experimental tests. However, the experimental data with fault are still correctly classified. The situation is similar for the data with fault at FV-2 (*H6*), shown as the black triangles. As far as the test data for the FV-1 fault and CV-5 fault (*H2*) are concerned, it can be seen that they all lead to decreased level L1 because of loss of inventory. FV-1 is clearly separated from the CV-5 fault because of the different patterns in temperature response. FV-1 simulates a pipe break right before the heater inlet. The actual fluid flow to cool the heater is reduced when FV-1 is opened. Consequently, the temperature difference across the heater will increase considerably. The loss of inventory through CV-5 does not reduce the flow through the heater. Therefore, the temperature difference is not affected. As to

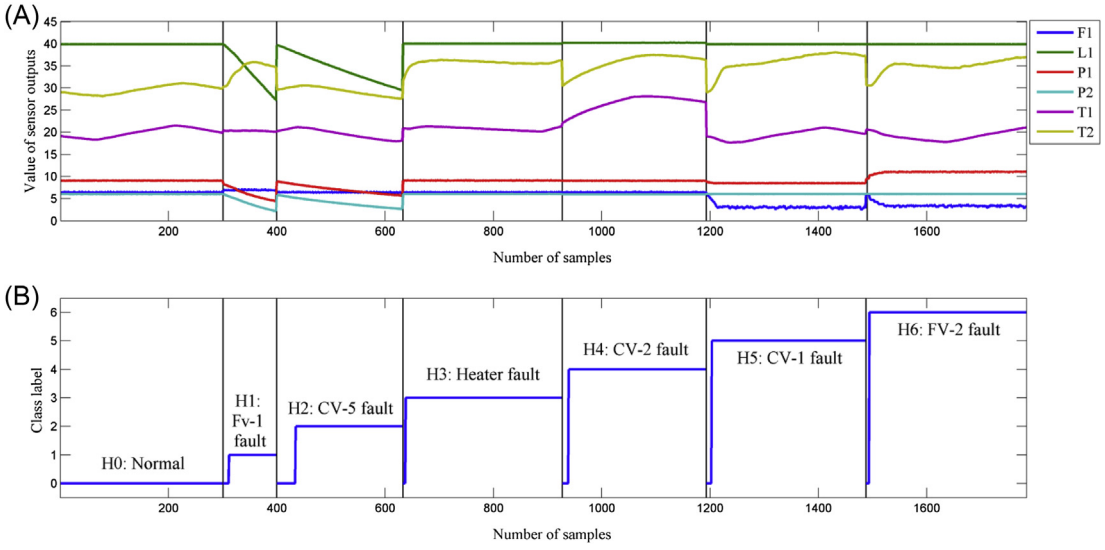

Fig. 8 – (A) Experiment data collected on nuclear power plant control test facility and (B) semisupervised classification results.

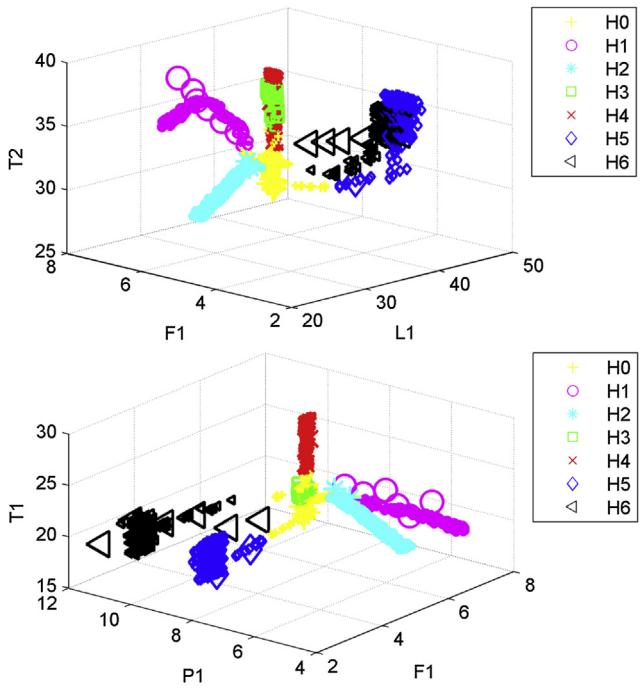

Fig. 9 – Test data, labeled data samples, and classification results.

the faults at CV-1 (*H5*) and FV-2 (*H6*), their effects on the three variables used in the first plot are similar. As a result, it can be seen that the two sets of data fall in the same region. The two faults are separated by the different effects on pressure P1, which is illustrated later in the second plot. The data with faults at the heater (*H3*) and CV-2 (*H4*) also overlap in the first plot, demonstrating the fact that those two faults also exhibit similar symptoms. However, it can be seen that they are correctly classified to different classes by the SSC model. The differences between the data with CV-1 fault (*H5*) and FV-2 fault (*H6*) become clear in the second plot at the bottom of Fig. 9, as the FV-2 fault will lead to increased coolant pressure P1, but the CV-1 fault is accompanied by slight decrease in P1. The differences between the two faults *H3* and *H4* are also clearer in the second plot. CV-2 close (*H4*) is the only fault that will lead to abnormal increase in T1.

In summary, satisfactory classification results of the six fault scenarios have been obtained using the SSC model. Considering the fact that the training data contain sizable uncertainties and labeled data are significantly smaller than that is typically required by a supervised model, the experimental tests demonstrated that the SSC-based fault diagnosis scheme proposed in this paper is a promising tool for diagnosis of process faults in a NPP.

4. Conclusion

A fault diagnosis scheme based on an SSC scheme is proposed in this paper for applications in NPP. Test results using plant data generated from a CANDU NPP plant training simulator as well as from a physical NPP test facility have shown that, even though only a few labeled data with uncertainties are used to train a classification model, all faults considered can be correctly diagnosed. The theoretical analysis and experimental validations demonstrate that the proposed scheme is a promising tool for fault diagnosis in NPP when it is normally difficult to obtain reliable training data under fault conditions for a supervised classification model.

Conflicts of interest

All contributing authors declare no conflicts of interest.

Acknowledgments

This research is financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Ottawa, Ontario, Canada and the University Network of Excellence in Nuclear Engineering (UNENE), Hamilton, Ontario, Canada. The authors would like to thank Ontario Power Generation (OPG), Toronto, Ontario, Canada for making the desktop CANDU NPP simulator available for this study.

REFERENCES

[1] C. Gottlieb, V. Arzhanov, W. Gudowski, N. Garis, Feasibility study on transient identification in nuclear power plants using support vector machines, Nucl. Technol. 155 (2006) 67–77.
[2] E. Zio, G. Gola, Neuro-fuzzy pattern classification for fault diagnosis in nuclear components, Ann. Nucl. Energy 33 (2006) 415–426.
[3] R. Razavi-Far, H. Davilu, V. Palade, C. Lucas, Model-based fault detection and isolation of a steam generator using neuro-fuzzy networks, Neurocomputing 72 (2009) 2939–2951.
[4] J. Ma, J. Jiang, Applications of fault detection and diagnosis methods in nuclear power plants: a review, Prog. Nucl. Energy 53 (2011) 255–266.
[5] J. Liu, R. Seraoui, V. Vitelli, E. Zio, Nuclear power plant components condition monitoring by probabilistic support vector machine, Ann. Nucl. Energy 56 (2011) 23–33.
[6] K. Moshkbar-Bakhshayesh, M.B. Ghofrani, Transient identification in nuclear power plants: a review, Prog. Nucl. Energy 67 (2013) 23–32.
[7] L.H. Chiang, M.E. Kotanchek, A.K. Kordon, Fault diagnosis based on fisher discriminant analysis and support vector machines, Comput. Chem. Eng. 28 (2004) 1389–1401.
[8] A. Widodo, B.S. Yang, Support vector machine in machine condition monitoring and fault diagnosis, Mech. Syst. Signal Process. 21 (2007) 2560–2574.
[9] S.K. Choi, J.Y. Park, I.B. Lee, Process monitoring using a gaussian mixture model via principal component analysis and discriminant analysis, Comput. Chem. Eng. (2004) 1377–1387.
[10] A. Widodo, B.S. Yang, Application of nonlinear feature extraction and support vector machines for fault diagnosis of induction motors, Expert Syst. Appl. 33 (2007) 241–250.
[11] Z. Zhu, Z. Song, A novel fault diagnosis system using pattern classification on kernel fda subspace, Expert Syst. Appl. 38 (2011) 6895–6905.
[12] P. Baraldi, F. Di Maio, Unsupervised clustering for fault diagnosis in nuclear power plant components, Int. J. Comput. Intell. Syst. 6 (2013) 764–777.
[13] J.H. Purba, A fuzzy-based reliability approach to evaluate basic events of fault tree analysis for nuclear power plant

probabilistic safety assessment, Ann. Nucl. Energy 70 (2014) 21–29.
[14] O. Chapelle, B. Schölkopf, A. Zien (Eds.), Semi-supervised Learning, The MIT Press, Cambridge, MA, 2006.
[15] P. Niyogi, Manifold regularization and semi-supervised learning: some theoretical analyses, J. Mach. Learn. Res. 14 (2013) 1229–1250.
[16] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., Wiley, 2000.
[17] E. Alpaydin, Introduction to Machine Learning, second ed., The MIT Press, Cambridge, MA, 2010.
[18] A.R. Webb, K.D. Copsey, Statistical Pattern Recognition, third ed., Wiley, 2011.
[19] X. Zhu, Semi-supervised Learning Literature Survey, Computer Sciences Technical Report 1530, University of Wisconsin, Madison, 2005.
[20] X. Zhu, A.B. Goldberg, Introduction to Semi-supervised Learning, Morgan and Claypool Publishers, San Rafael, 2009.
[21] M. Seeger, Learning with Labeled and Unlabeled Data, Technical Report No. EPFL-REPORT-161327, University of Edinburgh, Edinburgh, 2000.
[22] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
[23] A. Singh, R.D. Nowak, X. Zhu, Unlabeled data: now it helps, now it doesn't, Adv. Neural Inf. Process. Syst. 21 (2008) 1513–1520.
[24] T. Lu, Fundamental Limitations of Semi-supervised Learning, Master Thesis, University of Waterloo, Ontario, 2009.
[25] D. Zhou, O. Bousquet, T.N. Lal, J. Weston, B. Schölkopf, Learning with local and global consistency, Adv. Neural Inf. Process. Syst. 16 (2003) 321–328.
[26] G. Camps-Valls, T.V. Marsheva, D. Zhou, Semi-supervised graph-based hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 45 (2007) 3044–3054.
[27] O. Chapelle, J. Weston, B. Schölkopf, Cluster kernels for semi-supervised learning, Adv. Neural Inf. Process. Syst. 15 (2002) 585–592.
[28] J. Shi, J. Malik, Normalized cuts and image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2002) 888–905.
[29] A.Y. Ng, M.I. Jordan, Y. Weiss, On spectral clustering: analysis and an algorithm, Adv. Neural Inf. Process. Syst. 14 (2001) 849–856.
[30] X. Zhu, Semi-supervised Learning with Graphs, Ph.D. Thesis, Carnegie Mellon University, Pittsburgh, PA, 2005.
[31] F. Li, Dynamic Modeling, Sensor Placement Design, and Fault Diagnosis of Nuclear Desalination Systems, Ph.D. Thesis, University of Tennessee, Knoxville, 2011.
[32] M. Bhushan, R. Rengaswamy, Comprehensive design of a sensor network for chemical plants based on various diagnosability and reliability criteria. 1. Framew. Ind. Eng. Chem. Res. 41 (2002) 1826–1839.